

**TOWARDS A COMPUTATIONAL MODEL
OF ANAPHORA IN DISCOURSE:
REFERENCE TO EVENTS AND ACTIONS**

**Ethel Schuster
MS-CIS-86-34
LINC LAB 17**

**Department Of Computer and Information Science
School of Engineering and Applied Science
University of Pennsylvania
Philadelphia, PA 19104-6389**

(Revised June 1986)

Acknowledgements: This research was supported in part by DARPA grants N00014-85-K-0018 and N00014-85-K-0807, NSF grants DCR-86-07156, DCR 8501482, MCS8219196-CER, MCS-82-07294, 1 RO1-HL-29985-01, U.S. Army grants DAA6-29-84-K-0061, DAAB07-84-K-F077, U.S. Air Force grant 82-NM-299, AI Center grants NSF-MCS-83-05221, U.S. Army Research office grant ARO-DAA29-84-9-0027, Lord Corporation, RCA and Digital Equipment Corporation.

Towards a Computational Model of Anaphora in Discourse:

Reference to Events and Actions

Doctoral Thesis Proposal

9 June 1986

Ethel Schuster

Department of Computer & Information Science

The Moore School

University of Pennsylvania

Philadelphia, Pennsylvania 19104

Acknowledgements

I would like to thank my advisor, Bonnie Webber, for not giving up on me, for all her suggestions and examples, and for going over the numerous drafts that have lead to this one. I would also like to thank Dale Miller for his patience and his help in the formal aspects of this work, and Ellen Prince for answering all my questions. Many thanks go to Martha Pollack for her innumerable comments and help. Debbie Dahl, Julia Hirschberg, Megumi Kameyama, and Haim Levkowitz have read and criticized earlier versions of this proposal. I would also like to thank my committee members, Candy Sidner, Ellen Prince, Tony Kroch, Aravind Joshi, and Tim Finin for their comments and suggestions. Thanks also go to Richard Larson who spent many hours going through the literature with me, to my officemates, and to my friends for their support. Ve le-Eo: al hakol!

This work has been supported by the ARO grant DAAG29-84-K-0061.

Table of Contents

1. Introduction	2
1.1. Anaphora: How it benefits the Speaker	2
1.2. Special Problem of Anaphoric Reference to Events and Actions	4
2. Background	9
2.1. Studies on Anaphora	9
2.2. Anaphora and Discourse Models	14
2.2.1. Webber's Approach to Discourse Anaphora	14
2.2.2. Hans Kamp's Discourse Representation Theory	15
2.2.3. Model-Interpretive Anaphora	18
3. Formalizing Actions and Events in Discourse	20
3.1. Actions and Events: What are they?	20
3.1.1. Towards a Definition of Actions and Events	20
3.1.2. Vendler's Classification of Verbs	24
3.2. The Logical Structure of Action Sentences	27
3.2.1. Reichenbach's Work on Language	27
3.2.2. Davidson's Logical Form of Action Sentences	31
4. Reference to Actions and Events in Discourse	36
4.1. Discourse and Discourse Models	36
4.2. Formalizing the Representation	40
4.3. Generalizing the Representation: Entities evoked in the Discourse Model	45
5. Interpretation of Anaphoric References to Events and Actions	50
5.1. Issues affecting the Interpretation	50
5.1.1. Action and Event Focus	50
5.1.1.1. Focus in Discourse: Anaphoric References to Individuals	51
5.1.2. Additional Constraints	55
5.2. Mechanism at Work	56
6. Conclusions and Further Research	59

Towards a Computational Model of Anaphora in Discourse:

Reference to Events and Actions

Ethel Schuster

Abstract

When people talk or write, they refer to things, objects, events, actions, facts and/or states that have been mentioned before. Such context-dependent reference is called anaphora. In general, linguists and researchers working in artificial intelligence have looked at the problem of anaphora interpretation as that one of finding the correct antecedent for an anaphor—that is, the previous words or phrases to which the anaphor is linked. Lately, people working in the area of anaphora have suggested that in order for anaphors to be interpreted correctly, they must be interpreted by reference to entities evoked by the previous discourse rather than in terms of their antecedents. In this recent work, people have focused on entities of type concrete individual (an x) or set of such individuals (some x s) or generic class of such individuals (x s).

This proposal focuses on anaphora interpreted as referring to entities of type event and action. It considers four issues: (i) what aspects of the discourse give evidence of the events and actions the speaker is talking about, (ii) how actions and events are represented in the listener's discourse model, (iii) how to delimit the set of events and actions which correspond to possible choices for a particular anaphor, and (iv) how to obtain the speaker's intended referent to an action or event from that set of possible choices. Anaphoric forms that are used to refer to entities of type action and event include *sentential-it*, *sentential-that* pronominalizations as well as *do it*, *do that*, and *do this* forms. I will concentrate on the four previously mentioned issues along with other mechanisms that will provide us with better tools for the successful interpretation of anaphoric reference in discourse.

CHAPTER I

Introduction

1.1. Anaphora: How it benefits the Speaker

When people speak or write, they refer to things, objects, events, actions, facts and/or states that they (or other people) referred to before. They use certain words to "refer" or "point" to those things previously mentioned. Such context-dependent reference is called anaphora, and the words used to "refer" are called anaphors. Some examples of anaphors include pronouns such as *he*, *she*, *it* and noun phrases that contain *this*, *that*, *these* and *those*. Consider for example the following pair of sentences:

- (1) John was shot in broad daylight in Philadelphia. *He* was taken
to the University Hospital.
he = John

The *he* in the second sentence is the anaphor which points back to 'John' in the first sentence.

In the past, research in linguistics and psychology defined the problem of anaphora as that of finding the antecedent of an anaphoric expression, that is, finding the word or phrase to which the anaphor is linked. Lately, work in artificial intelligence (AI) and linguistics has shown that the problem of anaphora is not solely that of identifying the antecedent but how the referent¹ is found, based not only on the linguistic forms of the text but on a conceptual model of the discourse [Webber 78, Sag and Hankamer 84]. Most of this work has been directed at resolving definite anaphora (definite pronouns as in example (1) and definite noun phrases) when they refer to individuals and sets of individuals, verb phrase deletion (VPD), and one anaphora. The work presented here

¹I consider the antecedent to be the linguistic text which gives rise to an entity, as compared to a referent which is a non-linguistic entity.

will focus on another kind of anaphora: definite anaphors that refer to events and actions.²

When human beings communicate, we benefit from the use of anaphoric forms which make our interactions more comfortable and intelligible. If we were forced to describe everything explicitly, our discourse would become very complex, long and time-consuming. Moreover it would become confusing and lead to potential misunderstandings. Consider the following example which is similar to example (1) but in which a pronoun is not used in the second sentence:

- (2) The man who spoke at the anti-aparthied rally was shot in broad daylight in Philadelphia. *The man who spoke at the anti-aparthied rally* was taken to the University Hospital and *the man who spoke at the anti-aparthied rally* was immediately brought into the surgery room to have the bullets removed.

where instead of using 'he' I have used 'the man who spoke at the anti-apartheid rally.' We as listeners may find these sentences very odd and lacking coherence. We may start wondering whether the speaker is referring to more than one man who spoke at the rally. If the speaker is referring to the same man, why does s/he not use the pronoun 'he' to refer to the same person. If the speaker fails to use the pronoun, the listener may believe that s/he is meant to conclude something special from this unnatural usage. This conclusion may be unwarranted.

The use of anaphoric forms is very prevalent in discourse. In most cases, the anaphoric forms serve as linkages between the sentences. That is the difference, for

²In this work, I will not be concerned with non-pronominal anaphoric references to events and actions as shown in the following example suggested by Debbie Dahl,

- (i) John was shot in broad daylight in Philadelphia. These attacks make everyone become concerned about street violence.

While as a result of the work here, events describable as "John being shot in broad daylight in Philadelphia", "people being shot in broad daylight in Philadelphia", and so on will be available for pronominal reference, I will not deal with the inferential problem of matching "attacks" with one of the generic shooting events—i.e., I will not prescribe how to reason about the semantic relationship between "attack" and "shooting".

example, between (2) and (1). In (1), the use of the pronoun 'he' provides the listener with a pointer to 'John', hence establishing the linkage. In (2), there seems to be no linkage.

1.2. Special Problem of Anaphoric Reference to Events and Actions

Anaphoric language in English has been divided into two groups by Hankamer and Sag [Hankamer and Sag 76]: those that can only receive their interpretation from some "linguistic antecedent" and those that can be assigned an interpretation on the basis of "non-linguistic" information (or pragmatic environment). The former, which they call surface anaphora, includes VPD,³ gapping,⁴ sluicing,⁵ stripping,⁶ and *so* anaphora⁷. This kind of anaphor requires the presence of a "linguistic antecedent", and is claimed to arise at a relatively syntactic level via the application of deletion transformations. Consider the following example

Here the *0* indicates the part that has been "deleted" under a Verb Phrase Deletion rule as follows:

³The verb phrase (preceded by an auxiliary) in a sentence is deleted if and when the "logical form" representation of that verb phrase is identical to that of a syntactically permitted antecedent, as in example (3) [Sag 76]. VPD is characterized by the fact that the auxiliary survives the ellipsis.

⁴An ellipsis rule that applies in coordinate structures to delete all but two major constituents from the right conjunct under identity with corresponding parts of the left conjunct [Hankamer and Sag 76, p. 410]:

(ii) Ehrlichman duped Haldeman, and Nixon, Ehrlichman.

⁵An elliptical construction in which a sentence contains an embedded question consisting only of a *wh*-phrase [Levin 82, p. 590]:

(iii) Someone likes Janet, but I don't know who.

⁶A rule that deletes everything in a clause under identity with corresponding parts of a preceding clause, except for one constituent (and sometimes a clause-initial adverb or negative) [Hankamer and Sag 76, p. 409]:

(iv) Alan likes to play volleybal, but not Sandy.

(v) Gwendolyn smokes marijuana, but seldom in her own apartment.

⁷According to [Hankamer and Sag 76], *so* anaphora "is an anaphoric flag that turns up in certain constructions when an S or VP disappears" (p. 415):

(vi) If you have not yet changed your socks, please do so immediately.

- (3) Paul Cardello was shot by a member of the Colombo family
and Gianni Scorcese was *0* too.
0 = shot by a member of the Colombo family.

X - VP - Y - VP - Z
S.D.: 1 2 3 4 5
S.C.: 1 2 3 0 5 condition: 2 = 4

In this work I will not be concerned with this kind of anaphora, which has been studied elsewhere [Chomsky 65, Sag 76, Webber 78, Reinhart 83].

Anaphora interpreted pragmatically is called by Hankamer and Sag deep anaphora, and it includes *sentential-it* pronominalization,⁸ *one* anaphora, and *null complement anaphora* (NCA).⁹ An example of the latter is:

- (4) I told John to go to the store, but he refused.

Originally they defined this category as that in which the anaphoric elements had to make reference to either (i) the *deep syntactic structure* (as developed by [Chomsky 65]) of sentences in the discourse or (ii) nonlinguistic elements present in the context of the utterance [Hankamer and Sag 76]. They later suggested that the interpretation of a 'deep' anaphoric element is determined by reference to the *interpretation* of its antecedent (in all cases where there is one), i.e. by reference to some object in a model of the world constructed by the interpreter of the sentence of discourse [Sag and Hankamer 84, p. 328]. This means that the interpretation of 'deep anaphora'—now renamed *model-interpretive anaphora*—is not rigidly constrained by its antecedent in either surface or deep syntactic structure but by the relationship between the anaphor and an entity that forms part of the discourse model evoked by the discourse. The discourse model is basically a record of the content of the discourse. In it are represented entities corresponding to persons, to things, and to relations between those

⁸ [Sag 76] suggests that *do it* anaphora may be a special case of *sentential-it* anaphora (p. 243).

⁹ [Hankamer and Sag 76] argue that even though the NCA appears to be a syntactic deletion phenomenon, there seems to be no syntactic evidence that any such deletion occurs. For example, when the entire complement is missing, there is nothing left — unlike VPD where there is an auxiliary left — that shows that any syntactic operation took place within a real complement underlying the null complement.

entities, as well as to facts, events, actions, states, etc. This representation of a discourse model is relatively independent of the original sentence and its surface structure. I will follow this view in which the discourse model contains not only simple entities corresponding to things and people but to relations, properties and states as well as to actions and events.

The work I present here describes the process of dealing with anaphoric language when the reference is to events and actions.¹⁰ The overall goal is to make natural-language communication between human beings and machines more robust, and this can be done by providing computer systems with the ability to adequately interpret references to events and actions in discourse.

¹⁰M. Pollack has pointed out to me that events and actions may correspond to act-types and act-tokens in the philosophy literature. I will investigate these issues in the future.

Consider the following examples of reference to events and actions¹¹:

- (5) John was shot in Philadelphia. *It* happened at 10 am.
it = the just-mentioned event in which John was shot in Philadelphia.
- (6) John was shot in Philadelphia. *It* shocked me very much.
it = { (i) the event in which John was shot [in Philadelphia].
(ii) the fact that John was shot [in Philadelphia].
- (7) John was shot in Philadelphia. *It* has never happened before.
it = an event in which someone was shot [in Philadelphia].
- (8) John was shot in Philadelphia. *That* would never happen in Bogota.
that = an event in which someone was shot.
- (9) Al Pacino shot Mario Castellano yesterday. After he *did that*, he ran away in a red Cadillac.
that = shooting Mario Castellano yesterday.
- (10) John was shot by a Mafia member. Gianni Scorsese could not have *done it*.
it = shot someone/John.

In some of these examples (5, 6, 7, 8) the reference is to an event.¹² The differences among the sentences are based on whether the references are to the specific event: 'John being shot in Philadelphia', or to a different type of event: 'John being shot', 'Someone being shot'. In the remaining examples, the reference is to actions or types of actions. That is, I must address at least two specific issues: (i) what are the differences between a particular event or action and a type of event or action, *and* (ii) how much of the substructure of actions and events must be made explicit so as to

¹¹Parts of the interpretation of the anaphoric forms are in square brackets to indicate that they may or may be used in the interpretation. The reason for this will become clearer in the next section when I will discuss what forms part of an action. The first sentence makes all the information, including the one in square brackets, available. The question then, is how much of that information is used in the interpretation of the anaphoric pronoun.

¹²The reference in (6) could also be to a fact, as shown in the interpretation. It is not clear whether it will fall out naturally to consider references to facts, as well as to events and actions, in this thesis. Also, 'in Philadelphia' seems to be optional in the interpretation of the anaphoric expression, hence it appearing in square brackets.

accommodate references to them through the set of anaphora considered here. Given this, I must also address when the reference is to **specific event or action** and when, to a type. For example, the referent of *it* in (5) and (6) is the **specific event** of 'John being shot.' On the other hand, the referent of *it* in (7) and *that* in (8) is **any** event involving *somebody*(/John) being shot. And the referent in (10) is the action of shooting **somebody** or the action of shooting **John** while the referent in (9) is the action of shooting **Mario Castellano**.

This work will be concerned with the following issues:

- What aspects of the discourse give evidence of the events and actions the speaker is talking about,
- How are actions and events represented in the listener's discourse model.
It will also be necessary to have a representation that will allow us to identify the pieces, i.e. the substructure must be made explicit for future references of elements in it.
- How to identify the set of events and actions that could be possible choices.
- How to obtain the speaker's intended referent to an action or event from a set of possible choices. In what way, if any, can a focusing mechanism (such as has been proposed for interpreting anaphoric references to individuals) help in obtaining the speaker's intended referent.

CHAPTER II

Background

The problem of anaphora and the way in which a referent is chosen among several possible antecedents has been studied in artificial intelligence, linguistics and psychology. In artificial intelligence, the interest has sprouted from the desire to develop computer systems that communicate with their users using natural language and that are able to cope with anaphoric language in the same way we do. The concern in linguistics has focused mainly on what aspects of an anaphor-antecedent relationship are bound by syntax. Some work has also been done in trying to generalize this approach to antecedents in previous sentences in the discourse (e.g. [Reinhart 76] has looked at syntactic constraints on anaphora, [Reichman 78] has worked on discourse context, [Givon 83] presents a discourse framework with hierarchical structures that display discourse continuity, and [Schiffman 85] describes discourse constraints on *it* and *that* pronouns.) Psychologists have looked at anaphora from the point of view of memory and processing strategies.

2.1. Studies on Anaphora

In this section, I will review some of the existing approaches to anaphora. While there has been a lot of work done on anaphora, most of those studies have concentrated on interpreting anaphoric reference to individuals, objects, sets, and thing. Very few studies except [Robinson 81] have looked at other cases (e.g. anaphoric reference to states, facts, events, actions, etc.).

According to these studies, several things play a role in the correct interpretation of anaphoric referents. They include (i) syntactic and semantic constraints, (ii) inferences that the participants in the discourse make, and (iii) context in which the

anaphor occurs. Researchers in artificial intelligence have based heuristics on one or more of these aspects that allow them to find antecedents of particular anaphors. One example of the latter was Winograd's work [Winograd 72]. He developed a system—called SHRDLU—which could carry on a conversation with a human being about the states of a number of child's blocks located on top of a table. The main goal was for the program to obey orders concerning the blocks world and to answer questions about the current state as well as about past situations that came up in the conversations. The language used by the human user could involve a large number of expressions referring to the objects in the world such as 'the cube', 'a box', 'the green block', including quantified expressions such as 'all pyramids'. The program used a set of heuristics to find the antecedents. For instance, an expression like 'the cube' obtained its referent by examining previous sentences for mention of a cube. When the program encountered a pronoun, it used special heuristics which looked into the discourse for all of the different things they might refer to, and assigned a plausibility value to each interpretation based on factors such as its position in the syntactic structure, and the form of its determiner. Also, the program was able to handle simple references to events that took place during the conversation. In Winograd's words:

We have concentrated on the "local discourse context," and the ways in which parts of the meaning of a sentence can be referred to by elements of the next sentence. For example, pronouns like "it" and "they" can refer to objects which have been previously mentioned or to an entire event, as in "Why did you do *it*?" [...] and words like "that" can be used to mean "the one most recently mentioned," as in "Explain *that* sentence." (p. 157)

When a pronoun such as *it* was used in a sentence like

(11) Do it!

to refer to the entire main event of the previous sentence (which was saved in LASTEVENT), the program replaced the meaning of (11) with the description generated earlier for the event. When the pronoun was *that*, a different heuristic was used: *that* was taken to refer to the event most recently mentioned by anyone, while *it* refers to the event most recently mentioned by the speaker. One of the major criticisms of this work has been the fact that the heuristic approach lacked any kind of theoretical foundations.

The work on inference [Charniak 72, Rieger 74, Hobbs 76] has shown that in order to interpret anaphoric references, there is a need to make inferences based on world knowledge. These studies use forward and backward chaining methods of inference to find the antecedents. The approach is that anaphors are viewed as free variables which become bound using forward and backward chaining. The value that an anaphor gets bound to defines its interpretation. The problem with the inference mechanism is that there is no way to decide when to infer what and when to stop. Charniak, for example, used inference methods to find the antecedents of a pronoun in children's stories. His approach was to perform reference resolution on general text understanding tools, especially inferences needed for making sense of the situation provided by the text. Facts which might be relevant to the story were used to draw the inferences. If at a given time, all the facts that were necessary to make an inference were not available, they could 'wait' and look for the information until the inference would succeed. **Demons** were a type of fact which could assign a particular referent to a particular pronoun.

Hobbs [Hobbs 76] described two approaches to handling pronoun resolution during the analysis of English texts. The first depended only on the surface syntax of the sentences in the text. It was a naive algorithm that worked by traversing the surface parse trees of the sentences in a particular order looking for noun phrases whose gender and number matched those of the pronoun. In the second approach, pronoun resolution was handled during the semantic analysis of the text. The system used a data base of "world knowledge" inferences. It had four semantic operations which accessed the data base and the inferences were drawn selectively and in a context-dependent way in response to these operations. The first two operations tried to discover the relationship between sentences and to satisfy the demands made by predicates about their arguments. The third operation recognized and merged redundant expressions. Together, these three operations resolved most of the pronominal referents. The others, a fourth operation would try to resolve. This operation performed a bidirectional search of the text and "world knowledge" for an appropriate chain of inferences and used the naive algorithm.

Studies on syntax have had a different concern with respect to pronoun resolution: that is, what is allowed syntactically, not what is the actual referent. For example, Lasnik [Lasnik 76] proposed an interpretive framework in which pronouns are generated in deep structure and coreference between noun phrases is always possible except under a specific condition on noncoreference, which predicts the failure of coreference between certain noun phrases. He proposed the following noncoreference rule:

- (12) Given two noun phrases NP1, NP2 in a sentence, if NP1 precedes and commands¹³ NP2 and NP2 is not a pronoun, then NP1 and NP2 are noncoreferential.

That is, a pronoun and a noun phrase can corefer as long as the noncoreference rule does not apply.

Alternative descriptions of the noncoreference rule have been proposed by [Reinhart 76] who describes a c-command condition¹⁴, and Bresnan [Bresnan 82] who uses an f-command condition in Lexical Functional Grammar. Also, the Government and Binding theory of syntax (GB) proposed by Chomsky [Chomsky 81, Chomsky 82] develops the interpretive approach to anaphora. In general, all these syntactic approaches to anaphora have very little to say about the actual referent of a pronoun except for specifying which are the possible antecedents (within the given sentence) and which are not. In a typical syntactic approach, for example

- (13) *Jonathan* bought *him* an ice cream cone.

in (13), *Jonathan* and *him* are said not to corefer, but nothing is said about what *him* refers to. A pronoun may have a referent in an earlier sentence in the discourse or it may lack a referent altogether (e.g. the speaker may point to a male person nearby and indicate this person as being *him*). To address these concerns, other studies have focused on discourse approaches to anaphora, which I shall address next.

¹³NP1 precedes and commands NP2 when NP1 is to the left of NP2 and is not in a subordinate clause from which NP2 is excluded.

¹⁴This was later incorporated in Government and Binding Framework [Chomsky 81]

Grosz [Grosz 77, Grosz 81] described "focus spaces" as those parts of the speaker's knowledge which are relevant to the particular discourse segment. Whatever is in a focus space is highlighted via focusing. Several focus spaces may be relevant at one time but only one of them is available for processing at a given moment. Using this focusing mechanism along with the focus spaces, Grosz proposed a procedure for interpreting non-pronominal noun phrases. (This approach has been made more precise in [Grosz, Joshi & Weinstein 83].) Reichman [Reichman 78] expanded Grosz's mechanisms and described what she calls "context spaces" which are bound by their topics. Entities can have various focus levels within a context space and only those entities that are in high focus can be pronominalized. While this latter seems like a positive approach, there are still some problems in deciding what a topic is and how a listener can decide when an entity is in high focus or not.

[Sidner 79] described a mechanism for the correct interpretation—in English—of definite anaphora (e.g. pronouns and noun phrases used with *the*, *this* and *that*) in discourse. This mechanism, called (immediate) focusing, allows the participants in a conversation to center their attention on particular items of the discourse. When a speaker uses a definite noun phrase, s/he is signaling the listener towards the current item of the discourse which is in focus. A more detailed description of Sidner's focusing will be given in chapter V.

Lately, computational theories of overall discourse structure have been proposed [Grosz and Sidner 85] which, among other things, attempt to solve in a much cleaner way the problems Reichman noted.

Other studies have concentrated on describing the constraints on representations that limit the discourse entities and their corresponding discourse phrases which make the entire discourse coherent. In the following section, I will look at three studies that exploit the idea of a **discourse model** in order to characterize discourse coherence.

2.2. Anaphora and Discourse Models

A piece of discourse is a set of sentences or pieces of text that are spoken by one or more speakers. Usually the sentences in a discourse are connected in a way that makes them comprehensible and coherent. We use language (or sentences in the discourse) to describe certain situations to our listeners, and we try to communicate these situations using sentences in a discourse. We do so by attempting to get our listeners to construct an appropriate model: a **discourse model**.

There have been several proposals which suggest that discourse comprehension goes as follows: The speaker and the listener each build a model of the discourse from the incoming sentences. As the discourse proceeds, they integrate the new pieces into their discourse models [Webber 78, Kamp 84, Sag and Hankamer 84]. These studies are also concerned with how anaphoric expressions are related to their antecedents. By having the discourse models, the speaker and listener introduce "entities" in their discourse models as they are introduced into the discourse. When later on, these "entities" are referred to via an anaphoric expression, the discourse participants can make the appropriate link to that entity and hence are able to obtain the correct interpretation.

2.2.1. Webber's Approach to Discourse Anaphora

Webber [Webber 78] makes the above assumption. She considers a discourse model as made up of *discourse entities*: entities (which can be sets or individuals) "naturally evoked" by the discourse and linked together by the relations they participate in. For example, consider (14) followed by (15)

(14) John ate a banana.

(15) Then *he* got sick.

In this pair of sentences, the referent of the pronoun 'he' is some entity "naturally evoked" by the first sentence. In this case, it happens to be the entity evoked by the noun phrase 'John'. For a speaker, the referent of the pronoun is an entity in his/her

model which s/he expects the listener to have as well. Discourse entities can be evoked in the listener's model in several ways: (i) linguistically, from the text itself; (ii) perceptually or "pragmatically controlled", from the spatio-temporal environment; and (iii) inferentially, by reasoning from the existence of other discourse entities.

The types of discourse entities that Webber focuses on are primarily ones corresponding to individuals and sets, and she considers them in terms of definite pronoun anaphora and one-anaphora. She also proposes an analysis verb phrase deletion using some of the representational techniques proposed for the other two types of anaphora.

2.2.2. Hans Kamp's Discourse Representation Theory

Discourse Representation Theory is an approach to the semantics of natural language that relies on a level of analysis that acts between the syntactic representation of a sentence or a sequence of sentences and its truth-conditions. It tries to incorporate techniques that are used in model-theoretic semantics with accounts of pragmatic issues of language.¹⁵

The key component in this approach is a formal method for constructing representations of natural language discourse. That is, as sentences are processed, a formal representation of each sentence is developed: a *Discourse Representation Structure—DRS*—using a formal rule system of sentence syntax. The syntax does not account for scope-disambiguation and anaphoric links. That is, the link between an antecedent *A* and its pronoun only gets established at the level of representation construction. Kamp also claims that "deictic and anaphoric pronouns select their referents from certain sets of antecedently available entities" (p. 6). Presently, the theory focuses on the resolution of anaphoric ambiguities inter-sententially.

¹⁵ [Heim 82] presents a very similar approach to the one proposed by H. Kamp.

The key to handling anaphora within DRT is to introduce a set of elements of discourse syntax called 'discourse referents' or 'reference markers'. Reference markers are introduced by the DRS-construction rules as they apply to NPs. Their role in the account of anaphora is this: whenever two NPs are to be interpreted as anaphorically linked, a statement that equates the relevant referent markers is added to the DRS. Semantics are handled by a careful formulation of the truth definition for DRSs.

Consider for example the pair of sentences

(16) Pedro owns Chiquita. He beats her.

The DR of the first sentence will contain two elements, u and v which represent respectively, Pedro and Chiquita. Pedro has syntactically larger scope than Chiquita. The DR also contains the information that u owns v . This is represented as follows:

$m_1(16)$

	u	v
	.	.
	Pedro owns Chiquita	
	u = Pedro	
	v = Chiquita	
	u owns v	

where $m_1(16)$ is the label of the DR—structure—of the first sentence in (16). u and v are elements in the DR which represent, respectively, Pedro and Chiquita. The dots in the DR under these elements indicate that they are available or instantiated in the discourse.

The DRS-construction recipe requires an application of a rule for proper names. In this part (first sentence) the proper names 'Pedro' and 'Chiquita' are introduced as reference markers. Next, the 'u = Pedro' and 'v = Chiquita' are added to the set of statements in the box. 'u = Pedro' and 'v = Chiquita' are atomic sentences so no rules apply to them, and DRS-construction proceeds on.

Next, we have to incorporate the information from the second sentence of (16) by

extending $m_1(16)$. In order to do that, we have to decide on the referents of *he* and *her*. It seems natural that *Pedro* refers to *he* and *Chiquita* refers to *her*.¹⁶ So we expand $m_1(16)$ into:

$m(16)$

u	v
.	.
Pedro owns Chiquita	
u = Pedro	
v = Chiquita	
u owns v	
He beats her	
u beats her	
u beats v	

The process continues until all statements have been decomposed into atomic statements. The problem is that not all DRS rules can be applied mechanically. The construction of a DRS is supposed to fix anaphoric links, so every time a pronoun is encountered in the construction process, there is a certain margin for arbitrary choice. For example, the DRS for $m(16)$ establishes a link between *Pedro* and *he* but a DRS in which *he* is linked to an appropriate NP in previous discourse is also possible. One has only to replace the new 'reference marker', such as w in a statement like ' $w = u$ '.

If we assume that (16) was uttered by itself, without any previous discourse, we can easily justify the referents to be as in the example. Problems arise if (16) is uttered along with a deictic gesture or with some preceding discourse. These two situations are not handled in DRT. In general, DRT provides us with a nice structure for handling anaphoric links in discourse but it fails to go into the interpretation level.

¹⁶There is no explanation in the DR Theory that indicates how the referents are chosen. Assuming that the anaphoric links are understood when (16) was uttered by itself does not necessarily explain the choice of anaphoric links in the example above. Kamp claims that "*he* must refer to a male individual, and *her* to a female one" (p. 8). This, of course, is not so straight forward at all times, we don't always know whether certain antecedent is male or female. Kamp suggests that we ignore gender agreement (because it is not of primary concern to him), as well as other factors that help to disambiguate the reference of anaphoric pronouns.

2.2.3. Model-Interpretive Anaphora

In their 1984 work, Sag and Hankamer contend that there are two different anaphoric processes. One in which the "processes by which interpretations are assigned to elliptical verb phrases must be sensitive to scope of logical operators and variable binding" [Sag and Hankamer 84, p. 328].¹⁷ A second one wherein the "interpretation of a 'model-interpretive' anaphoric element is determined by reference to the *interpretation* of its antecedent (in all cases where there is one), i.e. by reference to some object in a model of the world constructed by the interpreter of the sentence of discourse" [Sag and Hankamer 84, p. 328]. This means that the interpretation of what they call *model-interpretive anaphora* is not necessarily based on the relationship between the anaphor and its antecedent but on the relationship between the anaphor and an entity that forms part of the discourse model evoked by the discourse.

Sag and Hankamer also assume that the speaker and the listener develop discourse models (containing 'entities' as in [Webber 78]). They not only contain 'entities' that represent individuals or sets of individuals such as those described by [Webber 78], but also entities that correspond to actions, events, facts, etc. For instance, in the examples in the introduction (5, 6, 7, 8, 10), the same sentence raises different entities that can be referred to subsequently.

Within this approach, both the speaker and the listener synthesize models of the discourse as they process the sentences that are introduced. These models are either linked to bigger models of the world or they are part of other models of the world. Then, the discourse becomes comprehensible by integrating the discourse that has been produced recently into the discourse model. For Sag and Hankamer—based on evidence provided by Johnson-Laird [Johnson-Laird 80]—a sentence is comprehended first as a propositional representation and, as the discourse proceeds, the propositional

¹⁷This one corresponds to what they earlier [Hankamer and Sag 76] called *surface* anaphora.

representation gets integrated into the discourse model. Eventually the propositional representations get deleted from the discourse model in order to make room for new ones. Given this approach, they assume that discourse understanding proceeds by producing logical forms and integrating them into the discourse model. They argue that this approach provides them with a way to explain the differences among the different kinds of anaphora, in particular, those that can be interpreted in terms of propositional representations and those that can be interpreted in terms of discourse models.

CHAPTER III

Formalizing Actions and Events in Discourse

3.1. Actions and Events: What are they?

I present here the description of what characterizes actions and/or events as the terms are used throughout this work. This description is based on the approaches presented in the linguistics and philosophy literature, some of which I will describe in this section.

3.1.1. Towards a Definition of Actions and Events

Philosophers, in particular, have been concerned with describing what characterizes actions. Many of them have characterized actions in terms of intentionality, i.e. whether an action is performed out of 'free will' or not. Davis [Davis 79] presents a review of the different theories that have tried to show what actions really are. Among the theories his review includes are: (i) the theory of agency, (ii) the contextual theory, and (iii) the volitional theory. Within the theory of agency, every action involves direct causation of an event by an agent. That is, there is causation by agents and causation by events which in turn may be caused by agents. Davis aligns himself with the volitional theory. According to him, actions and events can be defined as *doings*. That is, *doings* can be either actions or events and the differences between an action and any doings are based on whether there is 'deliberation' on the part of the agent to perform the doing or not. In other words, if a doing is deliberate, it is considered an action.¹⁸ If it is circumstantial to the agent, it is considered an event.

¹⁸ [Davis 79] describes actions in terms of 'volition', that is, actions require the desire or consciousness to be performed by the agent.

Also, for each type of action A , there is an event E_A but the opposite is not necessarily the case. One example Davis provides to describe actions is the following,

(17) Sam is running down the field; he stumbles.

If Sam had not been running, he would not have stumbled, i.e. his running 'caused' his stumbling. Is his stumbling an action? According to Davis it is usually not. In this example, 'stumbling' is considered the event-type which corresponds to "getting oneself to stumble." If Sam got himself to stumble, then it is considered an action because he acted as the agent of the action which was volitional. Another example is

(18) Sam coughed.

This carries the implication that the event in his respiratory track was caused by a volition. Here, Sam has performed an action, of the type 'getting oneself to cough'. If there was no volition, the case of Sam's coughing would be considered just a doing [Davis 79, p.26].

In the work presented here, I will not distinguish actions and events in this way because it does not appear significant with respect to their subsequent references. What is significant for me is that there are different anaphoric forms in English whose referents cannot be interchanged. I will say that one set refers to actions and the other refers to events. For instance, I will say *do it* refers to an action, as in the following example,

- (19)
- a. John runs everyday.
 - b. Mary *does it* every sunday.

where the *does it* refers to 'the action of running'. Now consider the case in which non-volitional actions such as 'bleed' or 'stumble' admit *do it* references. For example,

- (20) Rob de Castella stumbled at the end of the Boston Marathon.
He *did it* again when he received the winner's trophy.

In general, "non-action" verb phrases do not admit *do it* references as in (21)

- (21)
- a. John wanted to dance.
 - b. * Mary *did it* too.

while any verb phrase admits VPD as shown in (22):

- (22)
- a. John wanted to dance.
 - b. Mary did 0 too.

Here actions are taken as a sub-class of those things evoked (primarily) by verb phrases. Things one can access with *do it* are things an agent can do (or at least try!).

There is one set of sentences in which the action may be evoked indirectly by a verb phrase which doesn't itself describe an action. That is, there are sentences that can be referred to with *do it* and *do that* as if they were actions. Such are sentences that describe states¹⁹, that is, verbs such as 'like', 'hate', 'being', etc. Consider for example,

- (23)
- a. Moses Malone is 2 meters tall.
 - b. I can *do that* if I climb on two chairs.

where (23a) evokes a state, 'the state in which Mose Malone is 2 meters tall' which allows reference as if it were an action.

I will also take into account what the sentence evokes, i.e. the sentence evokes a particular event: the following example—(24)—evokes 'the event of John's throwing Mary's clothing into the garbage'.

- (24)
- a. John threw Mary's clothing into the garbage.
 - b. *It* made her cry.

Also, it will be important to note whether a sentence evokes one event or many. For example,

- (25) John told Mary to study everyday.
- a. *It* could help her pass her MCATS.
 - b. *It* made her angry.

(25) may allow one event accessible to pronominal reference or many.

These will be the lines that I will follow to characterize actions and events in this work. I will associate the events evoked by a sentence with its clauses and the actions

¹⁹These will be defined along Vendler's lines as shown in the next section.

with (some of) its predicates, e.g. in (19) the event corresponds to 'John running everyday' while the action is the 'running' predicate related to John. In general, events and actions are not interchangeable. For example:²⁰

(26) John lifted a piano. *It* bothered his mother, who thought he'd get a hernia.

$it = \begin{cases} \text{(i) the event in which John lifted a piano.} \\ \text{(ii) the fact that John lifted a piano.} \end{cases}$

in (26), the *it* can not be interpreted simply as the **action** of 'lifting the piano'. Lifting a piano does not bother John's mother: 'what bothers her is (the event or the fact of) John's lifting a piano: he's her son!'

In English, it is usually the case that actions are expressed by the verb phrase. I have not looked at other languages which do not have verb phrases and which still allow subsequent reference to actions and events.

The study of the different types of verbs goes back as far as Aristotle. Aristotle in the *Nicomachean Ethics* [Ostwald 62, Taylor 62] distinguished between two types of actions: (i) Movements, which take a specific time to bring about a state of affairs as a result of the action, and (ii) Operations, which terminate due to an agent's own doing and not to an external state of affairs. Operations in turn, were divided into Being and Becoming. The latter requires some activity on the part of the agent, and the former comes about due to some states of affairs obtained in the world.

Aristotle's classification of actions and event types serves as the foundation of much of the work on verb types in linguistics and philosophy. The first of these studies was done by Ryle [Ryle 49] who makes the distinction between achievements and activities. His achievements correspond to those events that culminate in some state of affairs while his activities denote events that can be done throughout some period of time and do not necessarily end at a specific time. Following Ryle, Kenny and Vendler

²⁰This example is due to Bonnie Webber

[Kenny 63, Vendler 67] developed verb classifications that are based on linguistic criteria. Vendler's work has been the most influential on later work by [Dowty 72, Taylor 77, Mourelatos 78, Steedman 82]. Because of its influence and importance in this work, I will focus on Vendler's work.

3.1.2. Vendler's Classification of Verbs

Vendler classified English verbs into four different categories: activities, accomplishments, achievements, and states, based on a temporal criterion: "the use of a verb may also suggest the particular way in which that verb presupposes and involves the notion of time" [Vendler 67, p. 97]. He defines the four categories by observing restrictions upon the co-occurrence of the verbs with time-adverbial phrases like 'in an hour', 'for an hour'. They are:

- *Activities*: call for periods of time that are not unique or definite. Examples include running, walking, pushing or pulling something.
- *Accomplishments*: imply the notion of unique and definite time periods. Writing a letter, making a chair, building a house, getting ready for something, and growing up are among this category.
- *Achievements*: involve unique and definite time instants. Recognizing, realizing, spotting and identifying something, losing or finding an object, crossing the border, are examples of achievements.
- *States*: involve time instants in an indefinite and nonunique sense. Verbs that describe states do not occur in the progressive,²¹ or in the imperative²² [Dowty 79]. Knowing or believing things, liking, loving, hating, and ruling are examples of this category.

Vendler uses two pairs of properties to classify the verbs: 'continuity' versus 'punctuality' and 'homogeneity' versus 'heterogeneity'. If a verb can occur in the

²¹For example,

(vii) * Mary is liking John.

²²For example,

(viii) * Like John.

progressive form, for instance, it is considered continuous. Also, if any part of the process is of the same nature as the whole, the verb is considered homogeneous. Using this criteria, the four categories are divided as follows:

	Activities	Accomplishments	Achievements	States
Continuous	YES	YES	NO	NO
Homogeneous	YES	NO	NO	YES

For Vendler, states and activities denote events whose subparts realize the same event. For example, if I say:

(27) I ran for an hour

it means that for any interval during that hour I ran. Then if I ran again for another hour, the sum of the processes represents a process of running. During any interval of those two hours, I was running. So states and activities refer homogeneously. That is, any part of the whole process is in itself a process of the same kind. Furthermore, if two processes of the same kind are combined, they form another process of the same kind. However there is a difference between states and activities: even though states may last over periods of time they "cannot be qualified as actions at all" [Vendler 67,p. 106].

Accomplishments and achievements, on the other hand, are heterogeneous. The difference between them is that the former have duration intrinsically while the latter can be placed within a period of time but cannot in themselves occur over or throughout a period of time. That is, achievements occur at a single moment. An example of accomplishments is

(28) Ann was drawing a circle at t .

means that t is on the time stretch in which Ann drew that circle. On the other hand an achievement is described by,

(29) Ann won a race between t_1 and t_2 .

and it means that the time instance at which Ann won the race is between t_1 and t_2 .

I believe that the difference between homogeneous and heterogeneous reference as introduced by Vendler—that is, the way in which an action or event as a whole stands

in relation to any subpart of that action or event—will play an important role in this work. For instance,

- (30) Haim promised to dance the samba non-stop for 10 hours.
 a. After he had been *doing it* for 2 hours, he fainted.
 b. He *did it* without the doctor's consent.

in (30) both actions (i.e. promising and dancing) are available for subsequent reference but (30a) requires an activity so it can only be 'dancing the samba...' while (30b) could refer to either.

There are some problems with Vendler's classification. Mourelatos [Mourelatos 78] has pointed out that it is not true that all achievements fail to occur in the progressive form. For example, Vendler considers the verb 'to die' as an achievement. This means that 'to die' could not occur in the progressive. But that is not the case in English where one can say things such as:

- (31) Jaffa was dying of cancer and *that* affected her husband's
 whole life
 $that = \begin{cases} (i) \text{ the event of Jaffa's dying of cancer.} \\ (ii) \text{ the fact that Jaffa was dying of cancer.} \end{cases}$

Some authors have noticed that very few verbs fit perfectly into each category. Most verbs can be made to take different categories depending on the particular adverb or auxiliary used, despite their basic category. Verkuyl [Verkuyl 72] argues that these categories should be seen as classifications of whole propositions instead of verbs. Steedman [Steedman 77] also argues that Vendler's categories can be grouped into two: one group includes activities, accomplishments, and achievements and they describe *events*: "what *happened* or what someone *did*" (p. 220). The other category does not describe events, it describes states and it belongs to a class of its own.

In summary, I believe the categorization of verbs as introduced by Vendler is an important part of the characterization of action and event descriptions and will certainly be needed in resolving action and event anaphora (i.e., choosing the intended referent).

Whether it will also be important in other aspects of this work (e.g., in generating additional needed action and event descriptions) is yet to be seen.

3.2. The Logical Structure of Action Sentences

Consider the following sentences:

(32) Haim ate a pepper

(33) What Haim did was to eat a pepper

According to logicians [Reichenbach 47, Davidson 67] sentences like (32) and (33) are action sentences. Davidson provides a description of the logical structure that underlies sentences such as (32) and (33), and much of his proposal is based on Reichenbach's ideas about the description of entities (individuals). Because of this, I will describe Reichenbach's work first and continue with Davidson's proposal about action sentences.

3.2.1. Reichenbach's Work on Language

Reichenbach, in trying to analyze how natural language has developed to characterize individuals, points out that the definition of the term 'individual' is a matter of convention. He denies the possibility of an individual being "something occupying a continuous and limited part of space and time" (p. 266). The reason for this, he claims, is that one may view a *pepper* as an individual and *vegetables* as the class of individuals to which the pepper belongs but, from a physicist's point of view, a pepper is composed of atoms in the same way vegetables are composed of peppers and cucumbers. Therefore, one has to forget about the "condition of physical connection of the parts and to consider the determination of the individual as a matter of convention" (p. 266-267). This view that the determination of individuals is a matter of convention, also applies to temporal individuals such as actions and events.

Reichenbach distinguishes between two types of individuals: (i) individuals of the

thing type: aggregates of matter keeping together for a certain time, and (ii) individuals of the event type: entities which are space-time coincidences and do not endure. He claims that an analysis shows that this same distinction between two types of individuals exists in natural language, and that language has developed forms of speech for both kinds of arguments. For example, a coronation, an assassination, an earthquake, are events, not things; but language contains designations of such events and uses them as arguments of sentences. Thus, a sentence like²³

(34) The coronation of George VI took place at Westminster Abbey.

describes a two-term relation between an event (the coronation event) and a thing (George VI) while

(35) The earthquake was followed by the explosion of the factory.

describes a two-term relation between two events (the earthquake and the explosion) . (Reichenbach does not comment on what 'Westminster Abbey' or 'the factory' are. I assume that they are *things*, as defined above.).

Reichenbach states that (34) is equivalent to

(36) George VI was crowned at Westminster Abbey.

which makes event-argument disappear and (35) can be transformed into

(37) The earth shook at the time t_1 and the factory exploded at the time t_2 , and $t_2 > t_1$.

In (37), the symbols t_1 and t_2 are event arguments, "since time points are events (or rather, classes of simultaneous events). There are some relations which can be formulated only as relations between events, such as time sequences; therefore event designations cannot be entirely eliminated in such sentences. Another important example of a relation between events is causality" (p. 268).

The claim that (34) and (35) can be transformed into (36) and (37) respectively has

²³These are Reichenbach's examples. For many people 'the coronation' may be a larger event (including parades, speeches, etc.) of which the actual 'crowning' event is only part.

been criticized by Davidson as being problematic because the sentences (34 and 36, and 35 and 37) have "quite different logical forms, but they are logically equivalent; one speaks literally of events while the other does not" [Davidson 67, p. 90]. This claim should not affect my work with respect to subsequent references. For instance,

(38) The coronation of George VI took place at Westminster Abbey. *It* happened at 10 am.

(39) George VI was crowned at Westminster Abbey. *It* happened at 10 am.

in both (38) and (39), *it* seems to refer to the same event: 'the event in which George VI was crowned at Westminster Abbey'.

Reichenbach uses the term situation to designate the object corresponding to a proposition. He describes a situation in a proposition as being composed of a function and argument which can be 'split' into argument-object and predicate-object (or property). Following the preceding argument —distinguishing individuals of the thing type and of the event type — there are two ways of splitting a situation: thing-splitting and event-splitting.

He claims, then, that an event and its property (or predicate) can be defined in terms of a thing and its property. So, one can have

$$g(v_1) \tag{3.1}$$

where v_1 denotes the event and g the predicate as being equivalent to

$$f(x_1) \tag{3.2}$$

"Thus, if $f(x_1)$ means 'George VI is crowned', g is the predicate 'coronation of George VI', which is a function of both the predicate 'is crowned' and the argument 'George VI'." [Reichenbach 47, p. 269].

Reichenbach then suggests using an asterisk to indicate the transition to event-splitting and write g in the form

$$[f(x_1)]^* \tag{3.3}$$

Then the expression $g(v_1)$ can be replaced by

$$[f(x_1)]^*(v_1) \tag{3.4}$$

The argument v_1 is the name of the event which has the property $[f(x_1)]^*$, and which is determined if both the predicate 'is crowned' and the argument 'George VI' are given. According to him, the event-argument v_1 can be written as²⁴

$$(\iota v)[f(x_1)]^*(v) \tag{3.5}$$

where the event is indicated by a bound variable v and where "this mode of expression, prevalent in conversational language, leads to the use of such predicates as 'takes place' and 'occurs', which merely express existence" (p. 269). Thus we say 'the coronation of George VI took place' which is represented in logical terms as

$$(\exists v)[f(x_1)]^*(v) \tag{3.6}$$

He goes on to suggest that "synonymously with the word *event* we shall use the word *fact*. The objective function $[f(x_1)]^*$ will be called a situational *fact-function*; ..." (p. 269).

In order to clarify this process, let us consider the sentence [Reichenbach 47, p. 270],

(40) Amundsen flew to the North Pole in May 1926.

which is represented in thing-splitting as:

$$f(x_1, y_1, t_1) \tag{3.7}$$

where f is 'flew' (or 'a flight'); x_1 is 'Amundsen'; y_1 is 'the North Pole' and t_1 is May 1926.

²⁴The notation using the iota-operator ι

$$(\iota x)f(x)$$

means 'the thing x having the property f ' [Reichenbach 47, p. 258]. For example,

(ix) London is the capital of England

is represented as

$$x_1 = (\iota x)c(x, z_1)$$

where x_1 means 'London', z_1 means 'England' and c means 'capital'.

This, in turn, can be transformed into event-splitting in various ways: One is using the fact-function $[f(x_1, y_1, t_1)]^*$,

$$(\exists v)[f(x_1, y_1, t_1)]^*(v) \quad (3.8)$$

and it is read a 'a flight by Amundsen to the North Pole in May 1926 took place' or 'there is an event which consists in the fact that Amundsen flew to the North Pole in May 1926'.

The second one is using the fact-function $[f(x_1, y_1)]^*$,

$$(\exists v)[f(x_1, y_1)]^*(v, t_1) \quad (3.9)$$

which is read as 'a flight by Amundsen to the North Pole took place in May 1926'. In this representation, an event-unit is considered to be 'a flight by Amundsen to the North Pole'. Many things can be said about this event-unit and it serves as a unit of quantification.

A third possibility is using the fact-function $[f(x_1)]^*$,

$$(\exists v)[f(x_1)]^*(v, y_1, t_1) \quad (3.10)$$

expressed as 'one of Amundsen's flights took place at the North Pole in May 1926'.

In summary, Reichenbach allows us to express a sentence such as (40) using 3 different fact-functions. These representations are logically equivalent but they have different logical forms to describe the same sentence. These representations do not seem to affect subsequent reference, that is, one can refer to the action(s) and/or event(s) evoked by the sentence, no matter what kind of representation is chosen.

3.2.2. Davidson's Logical Form of Action Sentences

Davidson [Davidson 67] starts by considering Reichenbach's approach to represent action sentences logically. That is, following Reichenbach one can transform (40) into

$$(\exists x)(x \text{ consists in the fact that Amundsen flew to the North Pole in May 1926}) \quad (3.11)$$

Davidson suggests that we treat that sentence in the following way

$(\exists x)(x \text{ consists in the fact that Amundsen flew to the North Pole and } x \text{ took place in May 1926})$

(3.12)

where the event is referred to explicitly as an object: events are introduced as "entities about which an indefinite number of things can be said" (p. 91). Davidson's argument for this analysis is that (i) it allows reference to events in discourse, and (ii) adverbs can be considered to be predicates of the event.

Davidson finds Reichenbach's approach to be attractive because

it eliminates a peculiar confusion that seemed to attach to the idea that sentences like (40) "describe an event." The difficulty was that one wavered between thinking of the sentence as describing or referring to that one flight Amundsen made in May 1926, or as describing a kind of event, or perhaps as describing (potentially?) several. As von Wright pointed out, any number of events might be described by a sentence like "Brutus kissed Caesar." This fog is dispelled in a way I find entirely persuasive in Reichenbach's proposal that ordinary action sentences have, in effect, an existential quantifier binding the action-variable. When we were tempted into thinking a sentence like (40) describes a single event we were misled: it does not describe any event at all. But if (40) is true, then there is an event which makes it true. (p. 91)

Davidson, however, objects to Reichenbach's analysis of action sentences. For one, the analysis may be applied to any sentence, no matter if it deals with events, actions, or any kind of statements. For example:

$$(41) 2 + 3 = 5$$

becomes

$$(\exists x)(x \text{ consists in the fact that } 2 + 3 = 5) \quad (3.13)$$

But one could continue recursively and the sentence could look like this

$$(\exists y)(y \text{ consist in the fact that } (\exists x)(x \text{ consists in the fact that } 2 + 3 = 5)) \quad (3.14)$$

and so on. I will focus on Davidson's approach that he claims improves on the problems discussed before.

Davidson starts with the basic notion that "verbs of action— verbs that say **what someone did**— should be construed as containing a place, for singular terms or variable, that they do not appear to"(p. 92). So, for a sentence like

$$(42) \text{ Shem kicked Shaun}$$

which normally consists in two names and a two-place predicate, he suggests that it be seen as a three-place predicate as follows:

$$(\exists x)(\text{Kicked}[\text{Shem}, \text{Shaun}, x]) \quad (3.15)$$

This sentence could be read as 'there is an event x such that x is a kicking of Shaun by Shem'. Davidson follows Reichenbach in describing the logical structure of action sentences in terms of events about which things can be predicated. His approach consists in two aspects: (i) it seeks to determine a basic structure of an action, and (ii) the predicate is analyzed as having a more complicated structure than the surface structure of the sentence displays. With respect to (i), he describes an action as

- "a relation between an agent, and an event, if the action is intransitive" or
- "a relation among an agent, a patient or accusative of the action, and an event, if the action is transitive, i.e. has an object." (p.105)

Going back to the example in (40), the event is taken to be 'Amundsen's flying to the North Pole' *and* 'Amundsen's flying to the North Pole in May 1926' as shown in (3.12). That is, Davidson analyzes the event argument as being a primitive.

With respect to (ii), Davidson suggests that what we must learn when "we learn the meaning of any predicate is how many places it has, and what sorts of entities the variables that hold these places range over. Some predicate have an event-place, some do not"(p. 93).

There are still several problems with the approach presented by Davidson. For instance, representing event sentences using explicit reference to the event as an object can become very complex, e.g. (3.12). Moore [Moore 81] has suggested introducing an event abstraction operator—EVABS—which will allow the introduction of event variables when needed:

$$(P X_1 \dots X_n) \leftrightarrow (\text{SOME } E (\text{EVENT } E) ((\text{EVABS } P) E X_1 \dots X_n))$$

So,

(43) John is kissing Mary

can be represented simply as

$$(\text{KISS JOHN MARY}) \quad (3.16)$$

whereas

(44) John is gently kissing Mary

requires the long form:

(SOME E (EVENT E) (AND ((EVABS KISS) E JOHN MARY) (GENTLE E))) (3.17)

To summarize, Davidson's approach proposes a treatment of action sentences in which events are considered to be individuals. There are several advantages that this approach provides:

(i) It simplifies the representation of sentences with adverbs of time and place (prepositions express relationships of their own and are separate from the verbs). For example:

- (45)
 a. Mary ate in my house.
 b. Mary ate on the bed.

describe an event of 'Mary eating' and its relationship with 'my house' or 'the bed'.

(ii) the sentence can be modified adverbially as in

(46) Mary ate quickly.

where 'quickly' describes an attribute of the action of Mary eating.

(iii) events can be objects of propositional attitudes as in:

(47) John didn't believe that Mary ate on the bed.

(iv) events can be arguments of sentences as in

(48) Mary ate in my house because she was hungry.

(v) events can be nominalized and can be referred to with pronouns as in

- (49)
 a. Mary ate and John watched it.
 b. Mary's eating made John hungry.

That is, for almost every sentence in discourse, there is an event that occurs that it refers to.

In this chapter, I have reviewed the literature on action and events. I have discussed what philosophers have had to say about what actions and event are. I have described Vendler's characterization of verbs and how this categorization affects the decision of what an action or a state is and whether the language describes when an action is completed or not. I have also reviewed Reichenbach's and Davidson's logical representation of action sentences. I will use the latter as a basis for my representation which I will describe in the following chapter.

CHAPTER IV

Reference to Actions and Events in Discourse

4.1. Discourse and Discourse Models

An overall concern of this work will be to describe how communication is done via discourse. A piece of discourse is a set of sentences or pieces of text that are spoken by one or more speakers. When we talk about discourse, we usually indicate the fact that those sentences in the discourse are connected in some way that makes them comprehensible and coherent. When we as speakers describe certain situations, we try to communicate these situations to our listeners. Speakers do so by attempting to get their listeners to construct an appropriate model: a **discourse model**.

The idea of a discourse model is not new [Webber 78, Sag and Hankamer 84] as described in the background chapter. A discourse model is viewed as containing representations of entities, along with their properties and relations they participate in. The key, then, in successful communication is for the speaker to transmit as much information about those representations of the entities, their properties and relations to the listener. It is usually the case that this cannot be done in a single utterance, so requires reference back to things mentioned previously. The speaker refers to a particular representation of a person, object, action or event in his/her discourse model and s/he can do so by using a pronoun, a definite noun phrase or a form such as *do-it*, *do-that* and so on. The speaker expects the listener (i) to have in his/her discourse model a representation for those objects, actions or events (or be able to create one with minimal effort), and (ii) to make the link between the referent and the representation of the entity using the rules of reference.

Discourse entities can represent individuals, sets, actions, events, facts, etc. The

referent of a definite pronoun or a sentential-it is then an entity in the speaker's discourse model which, given the goal of the communication, is expected to exist also in the listener's model. That is, an entity may be evoked by the discourse and included in the listener's discourse model in order to be referred to later on.

Given this view of discourse models which include discourse entities and their properties and relationships, my approach to anaphoric reference to actions and events will be as follows:

1.- As the discourse proceeds, each sentence is mapped onto a syntactic and semantic representation. The syntactic representation, maybe in the form of parse trees, should be able to provide me with information about the anaphoric expressions. For example, consider

(50) John saw him.

where we know that 'him' cannot refer to 'John' since if that were the case we would say

(51) John saw himself.

So, syntax does play a role in reference resolution.

Selectional restrictions on the verbs may also play a role in finding the referent of an anaphoric expression by specifying what kinds of things can serve as their subjects or objects. Consider

(52) Years ago they used to kill lambs. "Sacrifice" they called *it*.

(53) He had a little lamb. *It* had a white tail.

In both (52) and (53), the anaphoric expression is *it*. The question then, is how to recognize when an anaphoric expression such as *it* refers to an individual and when to an event. For instance, in (53) *he* must refer to a person and *it* must refer to an animal or a thing but not to a person.

2.- From the semantic representation of the sentence, I will derive a set of *entities*

evoked by that sentence in the discourse model. These discourse entities represent the individuals, sets, events, actions, and facts that have been introduced, which may be referred to by subsequent sentences. Note that while the primary focus of this work is on the reference to events and actions, I must include individuals, sets, facts, if they appear as discourse entities as well, in order to establish the relationship between the event entities and the individual discourse entities (such as 'John' which can be referred to with 'he') which may play a role in the correct interpretation of the discourse as well as in the correct characterization of the discourse model.

3.- When an anaphoric expression is encountered, it is checked against the set of available entities in the discourse model to discover which it draws its interpretation from. This may differ for *do it*, *do that* as well as for *it* and *that* pronominalizations. These anaphoric expressions may draw their interpretations from similar referents but may differ in how they change focus subsequently or in the conveyed 'position'²⁵ of the speaker. I shall try to adapt a focusing mechanism for events and actions similar to the those defined by [Grosz, Joshi & Weinstein 83, Sidner 83] for definite noun phrases. That is, relations between the event and action entities will be made explicit and a 'highlighting' between them (as used in focusing) will be used. So I will check whether the anaphoric expression refers to one of the events or a discourse entity (individual). If so, I have found the referent and the interpretation is completed.

4.- If the anaphoric expression does not refer to one of the evoked events and actions, methods I will develop will be applied for deriving additional discourse entities from the given ones (and, possibly, from properties ascribed to them). So far, I have identified two methods: generalization and a mapping of states to actions. Generalization basically consists in extending the specific set of events and actions evoked by the sentence to more general ones. Generalization states that a set A generalizes a set B if all the members of the set B are members of A. So, by

²⁵This is the notion of proximity.

generalizing I obtain additional action and event descriptions. For these derived descriptions, deciding whether an anaphoric expression refers to one results in the evocation of a new discourse entity with the more general description. For example,

(54) John was hit with a snowball last night in front of our house.
That happens all the time in our neighborhood.

in (54), the referent of the second sentence is not 'the event in John was hit with a snowball...' but 'an event in which a person gets hit with a snowball'. In this example, I have generalized from the specific event of John being hit with a snowball on a given night at a given place to the event in which people get hit with snowballs at any time in a given area of the city (i.e our neighborhood).

Note that the set of events and actions that are in focus can change as the generalizations are performed. That is, it may be the case that the most general event description available at a certain moment will be the one in focus. I will address the question of whether the focus is always the same or whether it moves accordingly as the generalizations occur.

Additional issues may be addressed when trying to obtain the correct interpretation of the anaphoric expression. It may important to check for the referent by using selectional restriction [Katz and Fodor 63] of the matrix. For example:

(55) She is *doing it* now.

may have as a referent an action (description) that appears in the progressive form while

(56) She usually *does it*.

may not have as a referent a specific action or an action that is completely instantiated.

5.- The second method that I have identified consists in mapping states to events for those event and action anaphors which cannot be resolved against entities evoked by the sentence. These procedures may involve some kind of inference mechanism that can derive events using the constraints specific in the sentence containing anaphoric terms.

An example of the latter may involve sentences like the following:

(57) Mary is a high school teacher. I could never *do that*.
do that = teach high school.

Here, the action evoked by the sentence is 'being a high school teacher' and the set of events includes all those in which Mary is a high school teacher. However, the preferred interpretation of the anaphoric expression *do that* in (57) is 'teach high school' rather than 'be a high school teacher'.

I will be looking for additional methods which may be needed to account for subsequent references to actions and events. I will also consider linguistic aspects such as the roles that relative clauses or embedded sentence play in anaphoric expressions.

So far I have outlined our approach to handling anaphoric references to events and actions. In the following section, I will describe the representation of actions and events evoked by the sentence that will be used throughout this work.

4.2. Formalizing the Representation

As described in the previous section, the first step in our approach is to map the sentences onto representations in order to indicate what they make available for future reference.

First, let us consider what characterizes actions or events. In most of cases, each clause²⁶ in the sentence (which may have one or more) gives rise to an event. As mentioned before, and as a way to establish the differences, I will (roughly) describe actions as being evoked by the verb phrase (the predicate of the sentence) and events as being evoked by the whole sentence. Events correspond to the conjunction of the action predicate with other predicates described in the sentence e.g. time, place, etc.

The auxiliary gives us a sense of whether the event is occurring (present), occurred (past), will occur (future) —or is just hypothesized— as well as whether an event is

²⁶A clause being defined as each S node in a sentence.

continuous or whether it has been completed.²⁷ These are all the issues described by [Vendler 67] that provide us with information about the event or the action. Events and actions usually have associated with them a time interval. This time interval can be present, past or future and it can have a well defined beginning and/or ending. There may also exist events which occur as points, that is, their beginning and ending time are the same as well as those in which there is no ending or no beginning. The verb phrase also establishes the relationship between the participants in the action. For example, in

(58) Mary cleaned the car.

the action of cleaning, which occurred in some past time was performed by Mary and it was done on the car. The 'clean' action relates Mary and the car. Tense and aspect usually indicate the sense of time of an action or an event. The tense indicates present, past or future and the aspect indicates the perfect (completion) and progressive (continuous) forms. In general, they are indicated syntactically by certain verbs and/or the auxiliaries and they play a role in the interpretation of the anaphoric reference to the actions and events.

Examples of tense and aspect markings are as follows:

(59)

- a. Mary cleans the car: present.
- b. Mary is cleaning the car: present, progressive.
- c. Mary cleaned the car: past.
- d. Mary had cleaned the car: past, perfect.
- e. Mary had been cleaning the car: past, perfect progressive.
- f. Mary will clean the car: future.
- g. Mary will be cleaning the car: future, progressive.
- h. Mary cleaned the car quickly: past.

Additional syntactic information within the sentence (e.g. adverbs) as in (59h) may

²⁷English has two divisions for aspect: (i) Progressive (verb *to be* and the verbal form *-ing*) and non-Progressive, and (ii) Perfect (verb *to have* and Past Participle) and non-Perfect. English also has a separate Habitual (which occurs only in the past tense), using the auxiliary form *used to*. The habitual can be replaced by its non-habitual equivalent, i.e. the non-habitual does not exclude habitual meaning [Comrie 76, p. 124].

also play a role in the interpretation of the anaphoric reference to the action. Hence, I need to characterize all the information that the sentence makes available to the discourse and introduce it in the representation of the sentences as they appear.

Assuming that a syntactic representation for each sentence is obtained, the next step is to describe the kind the semantic representation I will use for the sentences. One representation I am considering is a formal logical representation as described by [Miller and Nadathur 86]. I translate English sentences into logic based on Montague grammar (which can represent first-order logic within terms). This approach allows us to represent individuals in a logical and complex way. Remember the approach presented by Davidson [Davidson 67] in which events were considered individuals. In the representation that I will use, the logical form will describe the events that characterize the sentence (as compared to a proposition or individual), therefore allowing subsequent references to any one of the events evoked by the sentence. In this logical notation, I will have an 'extra' argument that is related to the main predicate. For example,

(60) Mary eats ice cream.

which is usually represented as

$$\text{eat}(\text{Mary}, \text{ice cream}) \quad (4.1)$$

will get mapped onto the following semantic representation

$$\exists v[\text{eat}(\text{Mary}, \text{ice cream}, v)] \quad (4.2)$$

which posits an individual eating event following [Davidson 67]. From this semantic representation, I will obtain the action evoked by the sentence which corresponds to its predicate (e.g. 'eat ice cream' in this example). This action can be represented in terms of λ predicates²⁸ as in,

$$\lambda x. [\text{eat}(x, \text{ice cream}, e)] \quad (4.3)$$

The sentence evokes an event description which will be represented as

²⁸This may change as the work described here progresses.

$$E: \lambda e. [\text{eat}(\text{Mary}, \text{ice cream}, e) \wedge \text{present}(e)] \quad (4.4)$$

meaning that **E** is an entity describable as 'the event in which Mary eats ice cream'. This is the entity associated with the event description evoked by the sentence. $\text{Present}(e)$ means that the event occurs now.²⁹ Following [Miller and Nadathur 86] I will use a λ notation that corresponds to a first-order predicate calculus, where the sentence is mapped onto a representation which corresponds to the event descriptions that it evokes.

As a way of illustrating the representation, consider an example given before:

(61) John was shot in broad daylight in Philadelphia.

which can be represented as follows:

$$E_{61-1}: \lambda e. [\exists x [\text{shoot}(x, \text{John}, e)] \wedge \exists t [\text{during}(t, e) \wedge t = \text{daylight_hours}] \wedge \text{in}(\text{Philadelphia}, e) \wedge \text{past}(e)] \quad (4.5)$$

Here, E_{61-1} can be described as 'an event in which John was shot in broad daylight in Philadelphia'. E_{61-1} corresponds to the real world event description in which John was shot and E_{61-1} took place at a time during the daylight_hours and E_{61-1} took place in Philadelphia. x is the person who did the shooting, who is not explicitly mentioned in the sentence. Also, $\text{past}(e)$ is defined as

$$\text{past}(e) \Rightarrow \text{time}(e) < \text{now}$$

So E_{61-1} makes a predication about the event in which John was shot and the details that form part of that predication. One interesting thing to note about this representation is that the event description must explicitly include an assertion that it has occurred, if it has.

Now consider another sentence,

²⁹ [Hobbs 85] describes a similar approach by introducing what he calls a "nominalization" operator ': "Corresponding to every n -ary predicate p there will be an $n + 1$ -ary predicate p' whose first argument can be thought of as the condition that holds when p is true of the subsequent arguments. Thus if $\text{run}(J)$ means that John runs, $\text{run}'(E, J)$ means that E is a running event by John, or John's running.The effect of this notational maneuver is to provide handles by which various predications can be grasped by higher predications (p. 62)".

(62) I told John to take out the garbage.

which describes two events: (i) the event of John taking out the garbage, and (ii) the event of the speaker telling John to take out the garbage. Note that only the latter has actually occurred while one can be non-committal about whether the former has. In this representation, both events must be accessible for future reference. That is, the representation must include the event of John taking out the garbage (E) and the event in which the speaker tells John to take out the garbage (E').

So, I start by describing the 'event of John taking out the garbage' which gets mapped onto the following representation:

$$E'_{62-1}: \lambda e'. [\text{take_out}(\text{John}, \text{garbage}, e')] \quad (4.6)$$

I can represent the other event description, E_{62-1} , as follows:

$$E_{62-1}: \lambda e. \exists v [E'_{62-1} v \wedge \text{tell}(I, \text{John}, v, e) \wedge \text{past}(e)] \quad (4.7)$$

The advantage of having this representation (as complicated as it may look) is that it allows us to generalize on various events. For instance, if we have the following sentence instead:

(63) I told John to shut up.

I can have the same representation for E_{62-1} (the event in which the speaker is telling John something) and I could change E'_{62-1} to be the description of the event in which John shuts up.

$$E'_{63-1}: \lambda e'. [\text{shut_up}(\text{John}, e')] \quad (4.8)$$

Note that in this example, $\text{past}(e)$ is defined as in the previous one (61), that is $\text{past}(e) \Rightarrow \text{time}(e) < \text{now}$

As another illustration, consider the following sentence:

(64) Julia was knitting during the colloquium.

which evokes an event in which Julia was knitting, represented as follows:

$$E_{64-1}: \lambda e. [\exists x [\text{knit}(\text{Julia}, x, e)] \wedge \text{during}(\text{colloquium}, e) \wedge \text{past}(e) \wedge \text{progressive}(e)] \quad (4.9)$$

where the representation is the entity describable as the event which Julia was knitting something (x) and E_{64-1} took place during the colloquium. Also, the event described by E_{64-1} took place in the past and the aspect of the sentence was progressive³⁰ which is defined as:

$\text{progressive}(e) \Rightarrow \text{aspect}(e) = \text{-ing, continuous}$

4.3. Generalizing the Representation: Entities evoked in the Discourse Model

We have available as an entity in the discourse model an event evoked by the sentence whose description I can generalize to other event descriptions. Generalization is defined as follows: If we have an event description E_1 and an event description E_2 , we say that E_1 generalizes E_2 if every event description denoted by E_2 is a member of the set of event descriptions denoted by E_1 .³¹ That is, if I want to see whether an event description E_2 is in E_1 , we can check to see whether it is in E_1 , i.e. if E_1 generalizes E_2 .

The reason for generalizing the description of an event evoked by the text is to allow it to be used in evoking another more general event which is the co-referent of an anaphoric expression.

As I generalize the representation of the event description that I have obtained for each sentence, I obtain more general representations of events. These event descriptions are available in the discourse model for future reference. The generalization of the events fits in an ordered structure in which the most specific representation (corresponding to the description of the event evoked by the original

³⁰I have chosen to indicate the progressive form with an explicit predicate in order to have that information available. This may change, that is, things like past, progressive, perfect may be included as properties of the events or as features of the event descriptions.

³¹This definition is based on Miller & Nadathur's [Miller and Nadathur 86] definition of *subsumption* (= generalization). They define it in terms of *concepts* as follows: "a concept C_1 subsumes another concept C_2 if every element of the set denoted by C_2 is a member of the set denoted by C_1 " (p. 6).

sentence) is at the bottom and the most general at the top. Since, within this approach, actions are part of the events evoked by the sentences, the generalization of actions will fall out of the generalization of events.

So if I consider (61)

(61) John was shot in broad daylight in Philadelphia.

and its representation E_{61-1} , I can generalize to 'any person y ' as follows:

$$E_{61-2}: \lambda e. \exists x \exists y [\text{shot}(x,y,e)] \wedge \exists t [\text{during}(t,e) \wedge t = \text{daylight_hours}] \wedge \text{in}(\text{Philadelphia},e) \wedge \text{past}(e) \quad (4.10)$$

We can further generalize to 'any place p ' and obtain:

$$E_{61-3}: \lambda e. \exists x \exists y [\text{shot}(x,y,e)] \wedge \exists t [\text{during}(t,e) \wedge t = \text{daylight_hours}] \wedge \exists p [\text{in}(p,e)] \wedge \text{past}(e) \quad (4.11)$$

And I can continue to generalize this last representation and obtain the following:

$$E_{61-4}: \lambda e. \exists x \exists y [\text{shot}(x,y,e)] \wedge \exists t [\text{during}(t,e) \wedge t = \text{daylight_hours}] \wedge \text{past}(e) \quad (4.12)$$

$$E_{61-5}: \lambda e. \exists x \exists y [\text{shot}(x,y,e)] \wedge \exists p [\text{in}(p,e)] \wedge \text{past}(e) \quad (4.13)$$

$$E_{61-6}: \lambda e. \exists x \exists y [\text{shot}(x,y,e)] \wedge \text{past}(e) \quad (4.14)$$

$$E_{61-7}: \lambda e. \exists x \exists y [\text{shot}(x,y,e)] \quad (4.15)$$

This set of generalizations can be ordered in a graph-like structure as shown in figure 4-1,

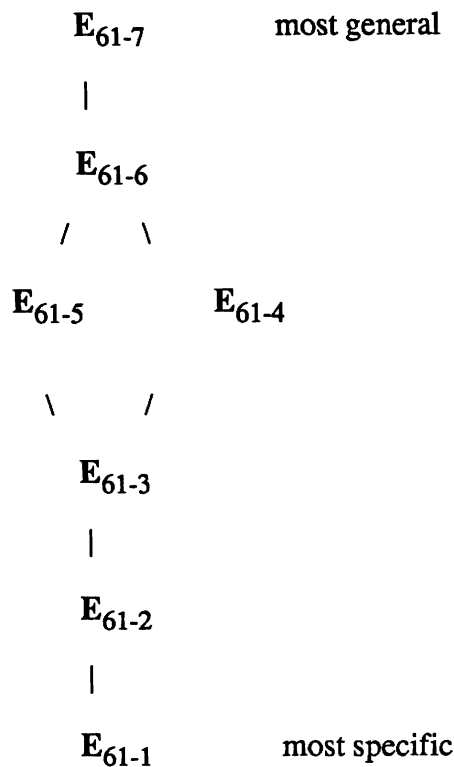


Figure 4-1: Generalization of 'John was shot in broad daylight in Philadelphia'

where I have gone from the most specific event description which is the one evoked by the sentence (bottom) to a more general one (top). Also, as specified by the generalization, every member of E_{61-1} is a member of E_{61-7} . This graph can help us to provide event descriptions that could fit into the sentences with the anaphoric expression.

Let us perform the same process of generalizing the representation for sentence (62)

I told John to take out the garbage.

given E'_{62-1} and E_{62-1} .

Let us start with the generalization of e' :

$$E'_{62-2}: \lambda e'. \exists s[\text{take-out}(s, \text{garbage}, e')] \quad (4.16)$$

to any person taking out the garbage and

$$E'_{62-3}: \lambda e'. \exists s \exists t [\text{take-out}(s, t, e')] \quad (4.17)$$

where any person can take out anything. The ordering of this event hierarchy is shown in figure 4-2.

E'_{62-3} more general

|

E'_{62-2}

|

E'_{62-1} most specific

Figure 4-2: Generalization of John takes out the garbage'

The generalizations of E_{62-1} include:

$$E_{62-2}: \lambda e. \exists v \exists s [E'_{62-1} v \wedge \text{tell}(s, \text{John}, v, e) \wedge \text{past}(e)] \quad (4.18)$$

where any speaker told John to take out the garbage.

$$E_{62-3}: \lambda e. \exists v \exists s \exists l [E'_{62-2} v \wedge \text{tell}(s, l, v, e) \wedge \text{past}(e)] \quad (4.19)$$

where any speaker told someone to take out the garbage.

The next generalization will allow for the event to occur at any time:

$$E_{62-4}: \lambda e. \exists v \exists s \exists l [E'_{62-2} v \wedge \text{tell}(s, l, v, e)] \quad (4.20)$$

while the last generalization will allow for the speaker to tell any listener anything (e.g. a joke)

$$E_{62-5}: \lambda e. \exists j \exists s \exists l [\text{tell}(s, l, j, e)] \quad (4.21)$$

which is the 'event of person x telling z to any person y .'

These representations correspond to the events evoked by sentence (62). Their ordering appears in figure 4-3,

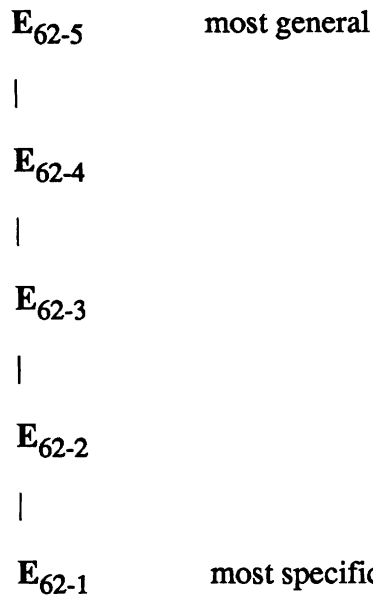


Figure 4-3: Generalization of 'I told John to take out the garbage'

where E_{62-5} is the most general set of events evoked and E_{62-1} is the most specific. One thing to notice is that in E_{62-3} the person being told about the garbage is the same and we could use a pronoun allowing for coreference as in 'I told John that he should take out the garbage' where 'he' and 'John' are coreferential. E_{62-4} does not require them to be coreferential.

So far, I have defined one possible representation for the events and actions evoked by the sentence. This representation allows me to create appropriate descriptions of actions and events evoked by the sentence which are available for future reference in the discourse. I will be looking into what other features are needed in an appropriate representation. I have shown how to generalize event descriptions, so that these more general event descriptions can in turn be associated with additional event entities required by the use of particular anaphoric expressions. Generalization of actions fall out of event generalizations. In the next chapter I will describe how the interpretation works and how the actions and events available from the sentence are used in this process.

CHAPTER V

Interpretation of Anaphoric References to Events and Actions

Recall that in the previous chapter I ended with a representation of a set of events entities evoked by the sentence and their descriptions. This set of events corresponds to the entities evoked by the sentence within the discourse. In this chapter, I will describe what kind of control mechanism can exist in order to correctly interpret the reference to the particular events evoked in the discourse. I will describe several constraints that affect this interpretation including the use of focus as introduced by [Grosz 77, Sidner 79, Sidner 81, Grosz, Joshi & Weinstein 83, Sidner 83]. I will also describe how to generalize the set of events in order to look for other potential referents of the anaphoric expressions.

5.1. Issues affecting the Interpretation

One of the issues that I want to address in this work is the reason for which we, as speakers, choose forms such as *do it* instead of *do that* and/or *that* instead of *it*. While there have been some theories that try to explain the differences [Halliday and Hasan 76], no real solutions have been provided. The most common explanation lies in the sense of proximity of the speaker to the object being referred to. That is, in general *it* and *this* seem to be 'closer' to the speaker while *that* is 'farther away' from the speaker.

5.1.1. Action and Event Focus

In this section I will describe what I shall call action and event focus. Action and event focus will be related to the focusing mechanisms proposed for interpreting anaphoric references to individuals. I will first describe in detail how focusing has been presented in terms of anaphoric referents to individuals and I will continue to show how this same mechanism could be extended to handle event focus.

5.1.1.1. Focus in Discourse: Anaphoric References to Individuals

Grosz [Grosz 77] first introduced the notion of focus in discourse, in which entities that are the most relevant at a given point become highlighted. They are used to handle implicit and explicit coreference. She proposed two kinds of focus: global focus which describes how large segments of the discourse relate to one another, and local focus which shows how individuals sentences 'connect' with one another to form discourse segments. The focusing model in general incorporates a representation of the knowledge of the language, the world, and of the discourse along with inferences and rules of linguistic performance.

Sidner describes a mechanism called immediate focusing (same as local focusing) in which she establishes the relationship between the immediate focus and definite anaphora. This relationship is as follows: (i) the items that are in the immediate focus are those that are talked about for part of the discourse, and (ii) the speaker uses definite anaphoric expressions to indicate his/her immediate focus [Sidner 81, p. 220]. She found that, in English, the distribution of pronouns is associated with the immediate focus while the distribution of other definite noun phrases is associated with global focusing.

The interpretation of a given definite anaphora depends on the shared focus that both the speaker and the listener have. Sidner describes an algorithm for finding the focus and for tracking it as the discourse proceeds. This algorithm has two parts given the local foci of a sentence: (i) it uses the local foci and a set of anaphor interpretation rules to interpret the anaphor in the following sentence, and (ii) it updates the foci, given the interpretation in (i). The rules to resolve anaphoric expressions are part of an inference mechanism. The focusing mechanism provides an ordered list of possible referents for a definite anaphor, which reduces the computational inferences in the search.

In general, a speaker directs the listener towards the pieces of discourse to look at

and the way to do so, that is, what to focus on. Objects (and for our purpose, actions and events) that come into focus are ordered in a hierarchy that indicates the degree of focusing, the highest one being the primary focus.

Sidner [Sidner 83] uses her focusing algorithm to provide an explanation for the differences in usage between *this* and *that* when used to refer to individuals. She contrasts the use of *this* with the use of *that*: when there are two objects of the same type, the speaker tends to use *this* for the most important one and *that* for the less important one. She then goes on to say that "in the case where the speaker wants to indicate that one of the two is more important, it will be co-specified using *this*; when *this* is used to mark relative importance, it will be referred to as primary focus" (p. 321).

Using several examples, she shows that when the speaker and the listener have different focus, *that* is used as the determiner of a definite noun phrase that co-specifies with the listener's focus while *this* is used with the speaker's focus. Also, when there are entities in the discourse that represent two different types of objects and one focus per speaker ("non-co-present", in her words), *this* specifies which element becomes the focus while *that* refers to those elements that stay out of focus.

Sidner's approach provides us with two important tools to handle the interpretation of anaphoric expressions which are: (i) a way to control the inferences performed by providing an ordered list of possible candidates as referents (a focusing algorithm), and (ii) allowing for certain nonidentity of reference (focus representation). This later point is captured in sentences in which different things can be referred to with either *it* or *they* as pointed out by Webber [Webber 78]. For example,

- (65) John gave Mary a banana.
 a. *It* was more than he gave me.
 b. *They* are loaded with potassium.

In Sidner's approach, *they* in (65b) can be correctly interpreted if the knowledge about "a banana", when activated, brings immediately its generic class into focus.

Several aspects of Sidner's immediate focus have been modified in [Grosz, Joshi & Weinstein 83]. They identified the center (or backward-looking-center— C_b) which is the "single entity that an individual utterance most centrally concerns [Grosz, Joshi & Weinstein 83, p. 44]." The center is related to the use of pronouns in English given the Centering Rule which says:

If the C_b of the current utterance is the same as the C_b of the previous utterance, a pronoun should be used.

They claim that this newer approach handles pronoun uses in a better way.

In general, the important contribution of the focusing approach (both [Sidner 83], and [Grosz 77] and [Grosz, Joshi & Weinstein 83]) is that it enables us to talk about non-linguistic entities. A center, as defined by [Grosz, Joshi & Weinstein 83] is in itself a non-linguistic entity. It is the semantic interpretation of a linguistic expression that may correspond to some concept in the mental representation of the speaker and the listener.

I propose to model my focusing approach to references to events and actions on the approach used to references to objects and individuals. What I will try to do in my work is investigate whether the focusing mechanism that handles references to individuals can, in its original form or in some modified form, work in handling references to events and actions.

Given the approach, I will try to indicate the differences between *do it* and *do that* as well as *it* and *that*. Consider the following cases: *VPD*, *do it*, and *do that*. *VPD* may indicate an "indefinite reference"³² and in the case of anaphora, it falls under the category of 'surface' (syntactically controlled).³³ For example:

(66) No one wanted to take the cat to the vet so John did *0*.
 0 = took the cat to the vet.

Now consider

³²"Indefinite" here is used in the same way that indefinite articles are used.

³³This is [Sag and Hankamer 84]'s classification: *VPD* is considered to be syntactically controlled.

(67) No one wanted to take the cat to the vet so John *did it*.
it = took the cat to the vet.

In this latter example, the reference is "definite"³⁴ to the particular action of 'taking the cat to the vet'. In (66) and (67), the referent of the intended action is the same. Focusing may help explain the differences between the two.

Now consider

(68) ?No one wanted to take the cat to the vet so John *did that*.
that = took the cat to the vet.

where the reference is "demonstrative"³⁵ and the sentence sounds very odd. So what does this show us? That I could consider the differences between "definite" and "demonstrative" references? Now, how about the focusing mechanism used? It may be that in the case of "definite reference", the action or event co-specifies with the (local) focus while in the case of "demonstrative reference" the action (or event) co-specifies with some already mentioned action but does not change the focus.

Now consider the following example,

(69) ...mostly I acted as a sounding board for Rollie. I've always
done that for him, and he's *done it* for me.

Here, it seems like *done that* brings the action of 'acting as a sounding board...' into focus, and once it is in focus, one can refer to it with *do it*. Notice how odd this example sounds with *it* and *that* reversed:

(70) ? ...mostly I acted as a sounding board for Rollie. I've always
done it for him, and he's *done that* for me.

The example sounds fine with both *do its*. When I replace both the *it* so that the sentence has both *do thats*, it seems that they both refer back to 'acting as a sounding board'. Then one sees they co-refer.

³⁴"Definite" here is used as in the article.

³⁵The demonstrative *that*.

A similar example is the following,

- (71)
- a. T.M. has been rescheduled to give the colloquium on the 19th of December.
 - b. I should be here for *that*, and I can't reschedule *it* again.

Here, *that* brings the 'colloquium' into focus while *it* refers to the 'colloquium' once it is in focus. Note how odd it would sound if I changed the second sentence into:

- (72) ? I should be here for *it*, and I can't reschedule *that* again.

I assume that the focusing mechanism will enhance the interpretation of the anaphoric expressions referring to events and actions. This mechanism can help us decide which of the event or action descriptions is the most highlighted, as it happens for references to individuals.

5.1.2. Additional Constraints

One tentative approach may be to consider additional information provided by each sentence. Each sentence may provide information that can constrain the search for the correct referent. That is, assuming I have the logical representation described above for each sentence and I encounter an anaphoric expression such as *it* or *that*, the next step is to find the correct referent for that anaphoric expression. For example, if I say

- (73) I ate three bags of Pepperidge Farm cookies. Haim *does it* everyday.

the first sentence makes available the following information to be filled in as arguments: **agent**: I; **patient**: Pepperidge Farm cookies; **amount**: three bags. The second sentence has an **agent**: Haim and a **when** argument corresponding to 'everyday'. In order to find the correct referent, the arguments specified in the second sentence overwrite those in the first sentence. Next, the arguments that are available from the first sentence characterize those that are empty in the second sentence. This information should provide enough detail as to what the referent is. If it fails, the next step is to try to fill the unspecified arguments in the second sentence with those of the next generalization of the first sentence (one higher in the generalization graph).

5.2. Mechanism at Work

Consider (61),

(61) John was shot in broad daylight in Philadelphia.

which I claimed before evokes the event entity describable as:

$$\mathbf{E}_{61-1}: \lambda e. \exists x[\text{shoot}(x, \text{John}, e)] \wedge \exists t [\text{during}(t, e) \wedge t = \text{daylight_hours}] \wedge \text{in}(\text{Philadelphia}, e) \wedge \text{past}(e) \quad (5.1)$$

If I consider the following sentence after (61),

(74) *It* happened at 7 am.

it = the just-mentioned event in which John was shot in Philadelphia.

the interpretation of the *sentential-it* will depend on how one associates the pronoun *it* with the event in which John was shot.

The question then is to find the referent from which the anaphor—*it*— draws its interpretation. Remember, I have the set of events that is available in the discourse model as referent for the anaphor. I can use focus, that is, I can adapt the ideas of focus to events and actions. *Do this, do that, do it, it* and *that* may use focus differently and, in turn, affect it differently.

So, I check to see whether there is an event or action that is the referent of the anaphor. If there is, I have found the referent and succeeded in the interpretation. If not, I can generalize the event description to try to tease out all the possible events evoked by the sentence. Once I have generalized the event descriptions and ordered them in a graph, I can choose among all those events and be guided by their order in the graph. For example, consider what the anaphor can refer to in each of the following sentences.

(75) Mary is cleaning the car. She usually *does that* on sundays.

In (75), *does that* cannot refer to a completely instantiated action since the sentence is making a generic statement while

(76) Mary cleans the car on sundays. She's *doing that* now.

in (76), *doing that* must refer to something that could appear in the continuous form. Alternatively,

- (77) Mary cleaned the car on Sunday.
 a. She *did it* for three hours.
 b. *It* took her three hours to *do it*.

where I can use the notion of achievement to justify the choice of referent.

Consider (62)

- (62) I told John to take out the garbage.

followed by

- (78)
 a. *It* is something that I need him to do.
 b. *It* is his responsibility.
 c. *That* shocked his mother.
 d. He had already *done it*.
 e. He was angry that I had *done that*.

As shown before, there are two events evoked in (62). In (78a) and (78b), the referent is the "embedded" event, 'John's taking out the garbage'. In (78c), *that* can refer to the "matrix" event: 'my telling John to take out the garbage' rather than 'John's taking out the garbage'. In (78d) the reference is to the "embedded" action of 'John taking out the garbage' while in (78e) it is the "matrix" action: 'my telling John to take out the garbage'. The issue here is whether the focusing mechanism can indeed help in the choosing of the embedded or the matrix referent.

In this section I have outlined a focusing mechanism for events and actions. This is based on the focusing approach described to handle anaphoric expressions referring to objects and individuals. As shown, the focusing mechanism provides an ordered list of possible referents as candidates. It describes the choice of anaphoric pronouns such as *this* and *that*. My hope is to extend this approach to handle referents to events and actions. Should the focusing mechanism fail to find a referent, a generalization approach is used along with it. That is, I can generalize the set of event entities available in the discourse model and use again focusing to choose among them. These will serve

as constraints that will affect the interpretation of the references to events and actions. I also hope that, in the same way that focusing reduces the inference mechanism in the search for anaphoric expressions when the referents are individuals, it will help reduce the inference when the referents are events and actions. There are probably many more issues that will come up with this approach which will enhance the interpretation further.

CHAPTER VI

Conclusions and Further Research

In this proposal I have outlined an approach for handling anaphoric references to actions and events in discourse. The main concern in the work will be the following: (i) to describe what aspects of the discourse give evidence of events and actions, (ii) to develop a representation for actions and events in sentences in order to refer to them appropriately, and (iii) to show how the given representation for actions and events facilitates the interpretation of the anaphoric forms when using a discourse model along with such constraints as focus, tense and aspect and so on.

The proposal started with a review of the literature on anaphora. Then I reviewed the literature on actions and events as developed by philosophers and linguists. Based on this, I tried to describe a possible representation for actions and events. My main concern is to find an adequate representation for events and actions so that they can be referred to in following discourse. After looking at some examples, many of the features that are necessary for the representation came up. For example, what is the minimal unit that we want to represent in a sentences such as (8)? Is it 'Shoot(x,z)', where x is 'the person doing the shooting' and z is 'the person being shot'? Now, the question is how to differentiate between a specific shooting and any shooting? How to encode this in the representation? Some of these questions have been answered and some have yet to be answered.

The next step I considered had to do with finding a representation for actions and events in discourse. I described the process by which each sentence is mapped onto a semantic representation and a set of event entities was obtained from that representation. I chose to represent events as λ expressions describing the entire predicate of the sentence, and actions as the predicate inside the events (i.e. verbs).

Given this representation, the set of events became available for future reference in the discourse. Both actions and events form part of discourse model and can be used to resolve anaphoric links. That is, actions and events represented in a formal way describe entities which are available for reference in the discourse model. When later on, a sentence appears which refers to the entity (event or the action evoked by a previous sentence) I are left with the need to establish the (anaphoric) link between the entity and its referent. I looked at the possibility of using a focusing mechanism to choose one referent over another. This mechanism could enable us to choose between the use of *do it* over the use of *do that* and viceversa, for example. The issue seems to be that *that* brings the action into focus while *it* considers the action that is in focus.

Most of the work in this proposal has described how I can handle references to events and actions. I have to describe a very precise representation to handle problems of continuity, parts of actions, and specificity of actions.

For the dissertation I will look very carefully at some data collected from various sources. The need to look at data carefully will help me identify the different pieces that will form part of the representation. The interpretation of the anaphoric links will be used in the discourse model together with the representation. I will look at the relationship between *it* and *that* and the roles they play in the reference to actions and events. I will also look at continued relatives clauses as in

(79) John left early which surprised his mother.

in order to see if the interpretive process is the same as in the cases outlined before.

Future research will focus on finding the a principled account for actions and events in discourse. I will try to refine the approach presented in the section dealing with the representation. I will also concentrate on developing and describing the constraints that affect the interpretation and the choice of the antecedent from the entities in the discourse model. I will try to develop one apparatus to see what is the degree to which reference to propositionally related entities such as facts, states, events, actions, etc. corresponds to the reference to objects.

I will also explore alternative representation formalisms that may give me a clue into the most adequate representation of actions and events for the purpose of reference.

Once this has been done, I will implement a computer system that handles this aspect of discourse. The main goal of the system will be to take a piece of the discourse, produce a representation for the actions and events, and produce the correct interpretation for the references used within that chunk of discourse.

References

- [Bresnan 82] Bresnan, J.
Control and Complementation.
In J. Bresnan (editor), *The Mental Representations of Grammatical Relations*, chapter 5, pages 282-390. MIT Press, Cambridge, MA, 1982.
- [Charniak 72] Charniak, E.
Toward a Model of Children's Story Comprehension.
Technical Report TR-266, MIT, 1972.
- [Chomsky 65] Chomsky, N.
Aspects of the Theory of Syntax.
MIT Press, Cambridge, MA, 1965.
- [Chomsky 81] Chomsky, N.
Lectures on Government and Binding.
Foris Publications, Dordrecht, 1981.
- [Chomsky 82] Chomsky, N.
Some Concepts and Consequences of the Theory of Government and Binding.
MIT Press, Cambridge, MA, 1982.
- [Comrie 76] Comrie, B.
Aspect.
Cambridge University Press, Cambridge, Great Britain, 1976.
- [Davidson 67] Davidson, D.
The Logical Form of Action Sentences.
In N. Rescher (editor), *The Logic of Decision and Action*, pages 81-95. University of Pittsburgh Press, Pittsburgh, PA, 1967.
- [Davis 79] Davis, L.
Foundations of Philosophy: Theory of Action.
Prentice-Hall, Inc., Englewood Cliff, NJ, 1979.
- [Dowty 72] Dowty, D.
Studies in the Logic of Verb Aspect and Time Reference in English.
PhD thesis, University of Texas at Austin, 1972.
- [Dowty 79] Dowty, D. R.
Word Meaning and Montague Grammar.
D. Reidel Publishing Co., Dordrecht, Holland, 1979.
Book with article about DO as part of the sentence structure.

- [Givon 83] Givon, T.
Topic Continuity in Discourse: A Quantitative Cross-language Study.
John Benjamins, Amsterdam, 1983.
- [Grosz 77] Grosz, B.
The Representation and Use of Focus in Dialogue Understanding.
Technical Note 151, SRI International, 1977.
- [Grosz 81] Grosz, B.
Focusing and Description in Natural Language Dialogues.
In A. Joshi, B. Webber and I. Sag (editor), *Elements of Discourse Understanding*, pages 84-105. Cambridge University Press,
Cambridge, England, 1981.
- [Grosz and Sidner 85] Grosz, B. and C. Sidner.
The Structures of Discourse Structure.
Technical Report 6097, BBN, November, 1985.
- [Grosz, Joshi & Weinstein 83] Grosz, B., A. Joshi and S. Weinstein.
Providing a Unified Account of Definite Noun Phrases in Discourse.
In *Proceedings of the Conference*, pages 44-50. ACL, Cambridge,
MA, June , 1983.
- [Halliday and Hasan 76] Halliday, M.A.K. and R. Hasan.
English Language: Cohesion in English.
Longman, London, England, 1976.
- [Hankamer and Sag 76] Hankamer, J. and I. Sag.
Deep and Surface Anaphora.
Linguistic Inquiry 7(3):391-426, Summer, 1976.
- [Heim 82] Heim, I.
The Semantics of Definite and Indefinite Noun Phrases.
PhD thesis, U. of Massachusetts, Amherst, September, 1982.
- [Hobbs 76] Hobbs, J.
Pronoun Resolution.
Research Report 76-1, City College, City University of New York,
1976.
- [Hobbs 85] Hobbs, J.
Ontological Promiscuity.
In *23rd Annual Meeting of the Association for Computational Linguistics*, pages 61-69. ACL, Chicago, Ill., July, 1985.

- [Johnson-Laird 80] Johnson-Laird, P.
Mental Models.
Stanford University, Palo Alto, CA, 1980.
- [Kamp 84] Kamp H.
A Theory of Truth and Semantic Representation.
In J. Groenendijk, T. M.V. Janssen and M. Stokoff (editor), *Truth, Interpretation and Information*, pages 1-41. Foris Publications, 1984.
- [Katz and Fodor 63] Katz, J. and J. Fodor.
The Structure of a Semantic Theory.
Language 39(2):170-210, 1963.
- [Kenny 63] Kenny, A.
Actions, Emotions and Will.
Humanities Press, New York, 1963.
- [Lasnik 76] Lasnik, H.
Remarks on Coreference.
Linguistic Analysis 2(1):1-22, 1976.
- [Levin 82] Levin, L.
Sluicing: A Lexical Interpretation Procedure.
In J. Bresnan (editor), *The Mental Representation of Grammatical Relations*, chapter 9, pages 590-654. MIT Press, Cambridge, MA, 1982.
- [Miller and Nadathur 86] Miller, D. and G. Nadathur.
Some Uses of Higher-Order Logic in Computational Linguistics.
In *Proceedings of ACL*. ACL, New York, July, 1986.
(to appear).
- [Moore 81] Moore, R.
Problems in Logical Form.
In *Proceedings of ACL*, pages 117-124. ACL, Stanford University, June, 1981.
- [Mourelatos 78] Mourelatos, Alexander.
Events, Processes and States.
Linguistics and Philosophy 2(3):415-434, 1978.
- [Ostwald 62] Ostwald.
Aristotle: Nicomachean Ethics.
The Bobbs-Merrill Company, Inc., New York, 1962.
Translated, with introduction and notes, by Martin Ostwald.

- [Reichenbach 47] Reichenbach H.
Elements of Symbolic Logic.
The Macmillan Company, New York, 1947.
- [Reichman 78] Reichman, R.
Conversational Coherency.
Cognitive Science 2(4):283-327, 1978.
- [Reinhart 76] Reinhart, T.
The Syntactic Domain of Anaphora.
PhD thesis, MIT, 1976.
- [Reinhart 83] Reinhart, T.
Coreference and Bound Anaphora: A Restatement of Anaphora
Questions.
Linguistics and Philosophy 6:47-88, 1983.
- [Rieger 74] Rieger, C.
*Conceptual Memory: A Theory and Computer Program for
Processing the Meaning Content of Natural Language
Utterances.*
Artificial Intelligence Lab Memo AIM-233, Stanford University,
1974.
- [Robinson 81] Robinson, A.
Determining Verb Phrase Referents in Dialogs.
American Journal of Computational Linguistics 7(1):1-16, January-
March, 1981.
- [Ryle 49] Ryle, Gilbert.
The Concept of Mind.
Barnes and Noble, Inc., New York, 1949.
- [Sag 76] Sag, I.
Deletion and Logical Form.
PhD thesis, MIT, June, 1976.
- [Sag and Hankamer 84] Sag, I. and J. Hankamer.
Toward a Theory of Anaphoric Processing.
Linguistic and Philosophy :325-345, 1984.
- [Schiffman 85] Schiffman, R.
*Discourse Constraints on 'it' and 'that' : A Study of Language Use in
Career-Counseling Interviews.*
PhD thesis, University of Chicago, June, 1985.

- [Sidner 79] Sidner, C.L.
Toward a Computational Theory of Definite Anaphora Comprehension in English.
 PhD thesis, MIT, June, 1979.
 Technical Report AI-TR-537.
- [Sidner 81] Sidner, C.
 Focusing for the Interpretation of Pronouns.
Journal of Computational Linguistics 7(4):217-231, 1981.
- [Sidner 83] Sidner, C. L.
 Focusing in the Comprehension of Definite Anaphora.
 In Michael Brady and Robert C. Berwick (editor), *Computational Models of Discourse*, pages 267-330. MIT Press, Cambridge, MA, 1983.
- [Steedman 77] Steedman, M.
 Verbs, Time and Modality.
Cognitive Science 1:216-234, 1977.
- [Steedman 82] Steedman, M.
 Reference to Past Time.
 In R.J. Jarvella and W. Klein (editor), *Speech, Place and Action*, pages 125-157. John Wiley and Sons Ltd., 1982.
- [Taylor 62] Taylor, A.E.
Aristotle on his Predecessors.
 The Open Court Publishing Company, Illinois, 1962.
 Translated from the text edition of W. Christ with introduction and notes.
- [Taylor 77] Taylor, B.
 Tense and Continuity.
Linguistics and Philosophy 1(2):199-220, 1977.
- [Vendler 67] Vendler, Z.
Linguistics in Philosophy.
 Cornell University Press, Ithaca, 1967.
- [Verkuyl 72] Verkuyl, H.J.
Foundations of Language: On the Compositional Nature of the Aspects.
 D. Reidel, Dordrecht, Holland, 1972.
- [Webber 78] Webber, B.L.
A Formal Approach to Discourse Anaphora.
 PhD thesis, Harvard University, May, 1978.

[Winograd 72] Winograd, T.
Understanding Natural Language.
Edinburgh University Press, Edinburgh, 1972.